VOL. 10, NO. 1 PHILO SPRING-SUMMER 2007

BIOLOGICAL FUNCTION AND NORMATIVITY

Kenneth G. Ferguson

Abstract: Ruth Millikan and others adopt a normative definition of biological functions that is heavily used in areas such as Millikan's teleosemantics, and also for emerging efforts to naturalize other areas of philosophy. I propose an experiment called the Lapse Test to determine exactly what form of normativity, if any, truly applies to biological functions. Millikan has not gone far enough in playing down as "impersonal" or "quasi" the precise mode of normativity that she attributes to biological functions. Further, her mode fails to qualify as genuine normativity at all, lacking an essential feature: some lapse of responsibility on the part of any entity or system that is charged with failing to do as it is "supposed." Nor, as we will see, is there anything in English idioms used to describe biological functions that can provide a persuasive argument to rehabilitate Millikan's normative definition.

Because biology is a natural science, we might reasonably expect that all of its attendant concepts could be defined naturalistically, without any need to import *normative* criteria. This would seem to include the concept of biological function, which might appear to be individuated sufficiently by the naturalized, *non*normative definition that

(1) F is the function of biological structure X if and only if F is what X does in the body.[1]

Yet there are influential philosophers of biology such as Larry Wright (1973), Ruth Millikan (1984, 1989), and Karen Neander (1991) who adopt a normative definition of biological functioning. According to this school of thought, the function F of any biological structure X is identified by what X has contributed to the inclusive fitness of ancestors from whom X was inherited (cf. Neander 1989, 173–174). Since it is the performance of F that has enabled the species with X to survive, these thinkers feel entitled to characterize functions in *normative* terms as

(2) F is the function of biological structure X if and only if F is what X *should* do in the body.

In Millikan's own specific language, "a proper function is a matter of . . . being 'supposed to' (impersonal) perform a certain function" (1984, 17). As Millikan adds that "the definition of proper function may also be read as a theoretical definition of 'purpose'" (1989, 291), we can characterize definition (2) as teleological.

On the teleological view, then, the function of biological structure X is what X is supposed to do in the biological system, its purpose, or as Wright has asserted, X's reason for existence (1973, 157). If a seemingly natural concept like biological function does truly harbor such surprising normative or teleological features, this is certainly something that philosophers could not afford to ignore. In fact, its repercussions for naturalized semantics (such as Millikan's own teleosemantics), naturalized ethics, and other drives toward the naturalization of areas of philosophy with biological underpinnings would be very serious. But this imposing theory must be tested.

1. NORMATIVITY AND RESPONSIBILITY: THE LAPSE TEST

One perplexing fact is that even those who most insist that biological functions be defined in a form semantically correspondent to (2) are nonetheless prone to deny that they are trading in normativity. Millikan (1984, 17), for example, commences her elaborate account of functions by cautioning the reader that her intended sense of "supposed to" is "nonnormative" (also "nonmysterious," she adds, apparently fearing that genuine normativity is bound to lead to something incomprehensible). Every so often, however, Millikan does slip in the word "quasi-normative" to describe her definition of functions (e.g., 1984, 5), and it is obvious, despite what cautionary notes she might add, that *some* manner of normativity, however thin, will be required if Millikan is to justify her entitlement to embed expressions such as "should" and "supposed to" into her definition. What we need to do, then, is to get a better grasp on exactly what is required to justify treating a structure or process as normative.

As a help in clarifying whether and in what possible sense biological functions may be normative, I propose the following simple test for recognizing normativity. First, we should note that function statements, as called for by supporters of the normative definition (2), are writable in the basic form typically used for the statement of normative claims, viz. O(F(x)), where 'O' is a sentence-operator indicating that "it should [or 'is supposed to'] be the case that ..." and 'F' a one-place predicate letter forming a sentence by taking a constant 'x.' Moreover, the pet example of the neo-teleologists, to wit, that the heart is for pumping blood (Wright 1973, 141), may presumably be cast, without remainder, into the form O(F(x)), as to read that

(3) it should be the case that the heart pumps blood.

Alternatively, they say that the *purpose* of the heart is pumping blood, that the heart is *for* pumping blood, that the heart exists *in order to* pump blood, etc. This talk about the heart may then be compared directly to a statement that is uncontroversially normative, say, for example,

(4) it should be the case that John contributes to the charity drive.

Both statements (3) and (4) are obviously intensional, since the modal operator creates an opaque context where the truth-value of O(F(x)) is not wholly determined by the truth-value of its embedded statement. This semantic circumstance suggests a simple test for normativity.

I shall maintain that for statements in the form O(F(x)) to count as genuinely normative in any sense, however thin, there must exist some kind of *lapse* or failure of responsibility in cases where 'O(F(x))' is true but 'F(x)' is false. It is simply incoherent to assert that a biological structure *should* do F but that it has no responsibility—at least in some tenuous, indirect, or metaphorical sense of the concept of responsibility—to see that F is done. Even though this failure of "responsibility" may not always be attributable to 'x' itself, especially in cases where F(x) is phrased in the passive voice, it is evident that there must be a lapse or failure of responsibility on the part of some entity or system if F(x) *should* or is supposed to be the case but isn't. This condition enables us not only to identify normative contexts but also to classify the mode of normativity in question according to the sort of lapse that is engendered. In (4), for example, *strong* or ordinary normativity is indicated given that if 'F(x)' is falsified under conditions where 'O(F(x))' is true, John's lapse is clearly *moral*, inciting blame or moral censure. The issue then devolves to the question as to exactly what sort of normativity, if any, is indicated by biological statements such as (3) that are so critical to the normativists' program. Once their projected form of normativity has been identified, we can then determine whether or not it passes our Lapse (of Responsibility) Test.

Using the opposable thumb as an example of a biological structure, we may apply our Lapse Test to a case captured by the following scheme of abbreviations:

(5) F(x): the opposable thumb grasps an object;
(6) O(F(x)): it should be the case that the opposable thumb grasps an object.

We are to think of a case where (6) is true but (5) is false, i.e., of a situation where the thumb *should* assist in grasping an object but does not. The question, then, is in what sense, if any, the thumb (x) may have lapsed in its "responsibility" to assist in grasping an object. And we are looking for at least some such sense that can justify the use of normative language in describing this state of affairs. Straightaway, we must clear aside the purely empirical sense of "should" occurring in statements such as "it should be the case that world temperatures will rise, because carbon dioxide is building up in the world's atmosphere." Although this equivocal sense of "should" has confused some who are writing in the area of evolutionary ethics, e.g. Robert Richards (see Ferguson 2001, 71–75), this alternative sense of the modality has no implications of normativity and is best thought of as synonymous with "it is predicted that . . ." For a genuinely normative sense of what the opposable thumb should do or is supposed to do, we surely must have in mind more than merely what the thumb is *predicted* to do, based on past cases, when confronted with a new opportunity for grasping. We must be careful to avoid allowing this usage of "should," when applied to biological structures, to affect our judgment about the reasonableness of statements such as that "the heart *should* do this or that," etc. This meaning of "should" has no normative tug and would not be adequate to achieve the goals of the neo-teleologists.

But neither is the relevant sense of "should" normative in the strong, basic sense already described above [as represented by statement (4)], because clearly no blame or moral censure is attributable to the thumb *as such* for its failure to grasp an object. Here, it is critical that we keep what the thumb itself should do separate from what a thumb-bearing animal should do *with* its thumbs; for example, it may be the case that a certain man should *use* his thumbs to assist in grasping a drowning child: the issue before us, however, is what the thumb itself should do. Such a distinction of course collapses if we confine our attention to organs such as hearts whose operation is not under the direction of the agent. For similar and equally obvious reasons, it can also be seen that the form of normativity attributed to the opposable thumb in statement (6) cannot be reasonably interpreted either as a matter of etiquette or of prudence, certainly not of rationality. As this exhausts the most familiar forms of normativity, our Lapse Test is still without a target.

Perhaps a good place to continue our search would be in the process of natural selection that normativists regard as the basis of the normativity found in biological functions. As Neander describes it, since what the opposable thumb contributed to the inclusive fitness of past generations of homo sapiens was its ability to assist in grasping objects, she concludes that grasping is therefore the stabilizing function (or purpose) of the thumb, to wit, what the opposable thumb *is for* or what it *should* do as part of the body. For this reason, its "function" is no longer merely what it actually does, but what it is supposed to do (Neander 1991, 168). If we generalize on this reasoning, we might arrive at the following way to capture Neander's supposed insight:

(7) x has performed function F in some entity E contained in world W at times $t_1, \ldots, t_n$ in the past.

(8) x's performance of F in world W at past times $t_1, \ldots, t_n$ has increased the inclusive fitness of entity E.

(9) Therefore, x *should* perform function F in entity E at some (appropriate) present or future time $> t_n$.

Assume that the entities in world W possess the properties required for evolution, viz., reproducibility, inheritability, a variation in traits, and adaptability to their environment, and that past times $t_1, \ldots, t_n$ cover more than one generation of the entity E. For lack of a better name, we might refer to the normativity articulated in (7)–(9) as Millikan-Neander normativity.

Not to be dismissive, but this reasoning suspiciously resembles a rather blunt instantiation of the Naturalistic Fallacy (cf. Hume 1990, 469–470). This can be seen more clearly if we divide Millikan-Neander normativity into two independent claims: first, steps (7) and (8), an empirical theory about how functions should be identified that we may call the Evolutionary Origin Thesis (EOT), and, second, step (9), an *inference* drawn from EOT, a normative conclusion about functions that we may call the Normative Conclusion (NC).[2] That NC does not follow from EOT hardly needs to be stated. Not only this, but this movement from a factual background to a normative result appears at odds with the way that evolution is ordinarily conceived, a story that certainly does not require any leap to assumptions about normativity. A useful digit, or some prototype from which such a digit later evolved, arises by *chance*, perhaps by mutation as is thought to be the case with the panda's "thumb" (Gould 1980, 22). Since the thumb incidentally is capable of effectively grasping objects because of its abductor muscles that enable it to pull away from the other fingers, it leaves animals possessing such a helpful digit more likely to survive; when animals with opposable thumbs do survive long enough to reproduce, this special apparatus is inherited by the next generation, with the opposable thumb eventually stabilizing as part of the biological makeup of the species. This account, as recognized from Aristotle onward (Physics Bk. II: Ch. 8 198b 21–32), can be understood completely without resorting to normative language.

Consequently, it does not appear that having a track record of resisting natural selection by availing itself of past opportunities for grasping can be the source of the thumb's present responsibility for grasping objects. There is simply nothing the thumb *should* do that bears any connotations of genuine normativity. To confirm this strong conclusion, we can switch our example from thumbs to a more "impersonal" natural structure such as *crystals* (cf. Bedau 1991, 650–654), "impersonal" being, in point of fact, an adjective that Millikan herself applies to her intended sense of normativity (1984, 17). Because they have all the characteristics already cited as sufficient for natural evolution, crystals are subject to the kind of natural selection that should be sufficient to produce outcomes of what we have dubbed Millikan-Neander normativity. But, plainly, we are not as tempted (as, perhaps, when discussing thumbs) to use normative language when describing the actions of inanimate crystals; there is definitely no cogent sense in which a crystalline component can be held *responsible* for failing to perform a function, no matter how often in the past the performance of this function has contributed to the inclusive fitness of its rock group. It appears, then, inappropriate to speak of what a crystalline component is "supposed to" do, even metaphorically. And the identical conclusion about geological structures can be transferred to the thumb, taking care, of course, to bracket this biological structure as a natural item without illicitly bringing in the interests or responsibilities of the animal *with* thumbs. The problem is not merely that the relevant sense of responsibility is very weak or thin in the case of these natural items: the problem is that there simply is *no sense of implied responsibility there at all*. As a consequence, Millikan-Neander normativity does not pass the Lapse Test, meaning that it does not seem necessary, or even really defensible, to define biological functions in normative terms.

2. NORMATIVITY AND ORDINARY-LANGUAGE ARGUMENTS

At this point, I am inclined to conclude that normative language has no business in a definition of biological functions. Also, I have found attempts by authors such as Millikan and Wright to promote a normative definition of functions by appealing to ordinary language to be unconvincing, even ironic. Ironic, because Millikan at least, claims to have no interest in what is "in the backs of people's minds when applying notions like 'in order to' and 'function'" (Millikan 1989, 290), dismissing ordinary-language searches for the precise meaning of philosophical terms as "a confused program, a philosophical chimera, a squaring of the circle, the misconceived child of a mistaken view of language and thought" (1989, 290). Her decidedly unenthusiastic attitude toward ordinary language might seem to imply that Millikan's own definition of biological functions is projected merely as a *stipulated* meaning[3] that would be totally indifferent to possible conflicts with English idioms. However, as supporters of the normative definition, including Millikan herself, often point to agreement with ordinary language as one of its assets, we are obliged to address their most noteworthy arguments dealing with ordinary language.

In short, concentrating on two of their arguments, neo-teleologists claim that a normative definition is needed in order to uphold (a) the ordinary-language distinction between functioning and malfunctioning and (b) the ordinary-language distinction between functions and mere effects. It is essential, before entering into our analysis of these two arguments, to remind ourselves that Millikan-Neander normativity actually consists of two components, an empirical component (EOT) and a normative component (NC), with the second being an (alleged) inference from the first. As a general response to both of these arguments from ordinary language, I will demonstrate that EOT by itself is adequate for drawing the normativists' desired distinctions, without any need to resort to NC. This will mean that even if the normativists have a means of accounting for ordinary-language distinctions, based on their grounding of functions in the process of evolution, still no implications follow as to whether we should define biological functions in *normative* terms. The propriety of this would have to be sustained by a proof of the validity of the entailment of NC from EOT, something we have already seen fit to challenge.

Turning to the first argument, due to Millikan, we encounter the claim that a normative definition of functions is necessary to account for the ordinary-language distinction between functioning and *mal*functioning. Granted, we often do tend to describe biological structures as malfunctioning whenever they are, appealing to Millikan's own list, "diseased, malformed, injured, broken, dysfunctional, etc." (1989, 295). The prefix "mal" does admittedly translate as "bad," and to Millikan it appears that a biological structure could not be correctly described as functioning *badly* unless it were "supposed to" perform a certain role as its proper function. After all, she asks, if there is nothing a heart or thumb (or linguistic device) is *supposed* to do, how can it go wrong? But surely it is possible to give a quite plausible rendering of the "mal" prefix simply by relying on EOT. In line with EOT, we can say that a biological structure X is correctly described as *mal*functioning when it is *not* functioning at a time when this structure is effectively needed to display a capacity which in the past contributed to the inclusive fitness of the ancestors from whom X was inherited. And notice that this strategy for drawing the distinction between "functioning" and "malfunctioning" does not require any resort to NC, the component of Millikan's theory that houses the normativity.

The fact that the English idiom favors the more colorful epithet *mal*functioning to describe such cases, as against merely *not* functioning (at needed times), can easily be explained by linguistic conventions occurring at the social level where the failure to function on the part of a biological structure (such as a lung or opposable thumb) may certainly be "bad" from the point of view of intentional, biologically-based agents with goals such as breathing and picking up objects that depend on the functioning of these structures. But any normativity contained in the English idiom surrounding the "mal" prefix can be completely explained by the interests of the agent: there is no reason to believe that it must exist as a biological primitive. This conclusion is even more obvious if we are willing to turn, as before, from hearts and fingers to Bedau's ingenious example of crystals (1991, 650–654), simply to remove the adulteration that the interests of the agent may have brought into the social development of this idiomatic distinction.

It appears, therefore, that we are able to explain the ordinary-language distinction between biological functioning and biological malfunctioning without being boxed into a normative definition of biological functions. The *non*normative component of the teleologists' definition (EOT) is by itself adequate for drawing the desired distinction, without any assumptions of normativity (NC). There is, however, another noteworthy argument derived from English idioms that supporters of the normative definition often launch. They claim that a normative definition of biological functions is needed to account for the ordinary-language distinction between the proper "functions" of a biological structure X and mere "effects" that X may produce in the body. According to Larry Wright, for example, a normative definition of functions is preferred in order to make sense of such plausible distinctions as "the function of the heart is pumping blood, not producing a thumping noise or making wiggly lines on electrocardiograms" (1973, 141). Clearly, however, the distinction that Wright has detected can easily be drawn by EOT alone, as was true with the first ordinary-language distinction we discussed. In line with EOT, we might say that it was the pumping of the blood that contributed to the past fitness of animals with hearts, not the pumping sound, certainly not effects hearts may display on modern hospital equipment designed to examine the organ for medical conditions. Functions, then, could be thought of as those activities of biological structures that are believed to have played relatively more significant roles in enhancing the evolutionary success of a species. But this strategy constitutes merely a productive means to categorize the effects of the structure, using EOT as a background theory. But since the entire classification can be done by EOT alone, that is, by the *non*normative part of the teleologists' definition of biological functions, the success of this classification shows absolutely nothing about the necessity of defining functions in *normative* terms.

Consequently, the dovetailing of the Millikan-Neander definition of functions with certain distinctions found in ordinary language does nothing to show that functions must be characterized normatively. EOT alone provides the agreement. Although this establishes my point vis-à-vis the alleged ordinary-language proof for the normativity of functions, I may also point out that the fact that EOT is *consistent* with certain ordinary-language distinctions does not necessarily mean that even EOT is needed to provide an adequate definition of biological functions. Many people have questioned whether biological functions must be defined in evolutionary terms, especially given that leading scientists such as William Harvey (1941, 49) made free use of the term "function" in the days before biology was generally given an evolutionary interpretation.[4] Moreover, it does appear that a biological structure may possibly have functions even if the creature with the structure *has not evolved*. Consider, for example, possible creatures whose conditions have always been static because of a lack of variation in traits: it would be implausible to maintain that the biological structures of such creatures therefore cannot have functions. Incidentally, Definition (1) itself is amenable to some further ordering among the various functions of biological structures, if such a pecking order is considered desirable for some scientific or other practical purpose. But such a split into "functions" and *mere* "effects" (or lesser functions) could be based on present conditions and present human interests, need not assume EOT, and certainly need not assume that this division is pre-sorted by normative conditions existing at the biological level.

Even though its apparent agreement with such distinctions of ordinary language as (a) and (b) may appear to be one of its strong suits, I will conclude by pointing out one result from the normative definition of functions that appears to clash with perceptions of ordinary-language usage. As we know, the normativists seek to highlight one (or at least a very limited subset) of the actions of a biological structure as its function(s), the rest being only effects. When Wright, for example, asserts that he takes "the definite article formulation [of function] as paradigmatic" (1973, 141), he announces his intention to deny to most of the actions of a biological structure (viz. their sonorous properties, effects on monitoring equipment, color displays, etc.) the honorific title of function. Only those that qualify as proper functions are then treated normatively, meaning that only statement (10) below would be treated as true based on the normative definition, with statements (11) and (12) treated as *false*:

(10) The heart should pump blood.
(11) The heart should produce a thumping noise.
(12) The heart should make wiggly lines on electrocardiograms.

Although this truth-value distribution makes some technical sense inside the normativist account of functions, wherein a biological structure is "supposed to" perform only those actions that count among its proper functions, it would be difficult to find any ordinary-language sense of "should" that would recognize such a mixed truth-value distribution as plausible. The problem is that any teleological sympathy that might be summoned to make (10) appear plausibly true would likely also carry the truth of (11) and (12) in its wake. Although this point of conflict with ordinary language is certainly not by itself enough reason to reject the normative definition of functions, it is yet another factor that weighs in the balance of considerations against definition (2). The meaning of "should" appearing in definition (2) simply does not have much resonance with ordinary language and, in any case, this modality is not needed for drawing distinctions (a) and (b), a burden fully met by the nonnormative hypothesis EOT alone.

3. CONCLUSION

If we are correct that the normative character of functions has not been satisfactorily established, Millikan's ambitious and impressive attempt to provide a naturalization of semantics is stymied. According to Millikan's so-called teleosemantics, a *true* sentence is one that maps onto the world as it is "supposed to," a false sentence one that fails to map as it is "supposed to" or should map (1984, 88), with these mapping relations allegedly grounded in the evolutionary history of the referring terms occurring in these sentences. By her own analogy, sentences are supposed to map onto the world just as "the color pattern of the chameleon is 'supposed' to match what it sits on" (1984, 89). But stripped of their normativity, Millikan's mapping relations simply collapse into those encountered in the classic correspondence theory of truth, a troubled semantic theory that Millikan had been hoping to regenerate with an infusion of normativity (1984, 9, 86–87), perhaps normativity artfully minimized as impersonal, quasi, derived, thin, or weak but still expected to determine the truth values of sentences.

Beyond its implications for Millikan's own special semantic project, the successful extinction of a primitive kind of normativity existing at the level of biological functions would clearly also impede ongoing efforts to naturalize or "biologize" other area of philosophy, especially in ethics (cf. Wilson 1975, 562). For it would effectively close one promising route to the long-sought Northwest Passage between biological facts and values that these philosophers have been seeking.

East Carolina University

ACKNOWLEDGEMENTS

Nick Georgalis, Frank Murphy, John Collins, Richard Miller, Mike Veber, and Paul Woodward all provided valuable criticism.

NOTES

1. The closest to (1) in the literature is the influential definition of functions by Robert Cummins, although I have elected not to include as part of the definition his condition that a function "is singled out by its role in an analysis of some capacity of the body," (Cummins, 1975, 165) as this added condition implies that biological functions are in some sense artifactitious, emerging only in the course of scientific investigation.

2. As can be seen in the section below, defenders of the normative definition of functions often bundle EOT and NC in a way that disguises their independence, drawing illicit sustenance from its more plausible component EOT in order to make NC, a problematic inference, more palatable.

3. Actually, Millikan denies that her definition of function should be considered merely stipulative, even though she says that this makes no material difference to her argument. Instead, she views it as "theoretical" (1989, 289-291); exactly what note, if any, a "theoretical" definition would be obligated to take of ordinary language would be a challenging question to answer.

4. Neander dismisses historical usage of the term "function" as irrelevant, saying, for example, that "functions might happen to be effects for which traits were selected by natural selection, whether or not Harvey or anyone else was aware of the fact" (1991, 176).

BIBLIOGRAPHY

Bedau, Mark. (1991). "Can Biological Teleology Be Naturalized?" *Journal of Philosophy* 88: 647–655.

Brandom, Robert (2001). "Modality, Normativity, and Intentionality." *Philosophy and Phenomenological Research* 63: 587–609.

Cummins, Robert. (1975). "Functional Analysis." *Philosophical Review* 82: 139–168.

Ferguson, K. G. (2001). "Semantic and Structural Problems in Evolutionary Ethics." *Biology and Philosophy* 16: 69–84.

Gould, Stephen J. (1980). *The Panda's Thumb*. New York: W. W. Norton.

Harvey, William. (1941). *Anatomical Studies on the Motion of the Heart and Blood*, 3rd ed., trans. Chauncey D. Leake. Springfield, Ill. Charles C. Thomas.

Hume, David. (1990). *A Treatise of Human Nature*. Oxford: Clarendon Press.

Millikan, Ruth. (1984). *Language, Thought, and other Biological Categories*. Cambridge, MA: MIT Press.

Millikan, Ruth (1989). "In Defense of Proper Functions." *Philosophy of Science* 56: 288–302.

Neander, Karen. (1991). "Functions as Selected Effects: The Conceptual Analyst's Defense." *Philosophy of Science* 58: 168–184.

Post, John. (1987) *The Faces of Existence: An Essay in Nonreductive Metaphysics*. Ithaca, N.Y.: Cornell University Press.

Rosen, Gideon. (2001). "Brandom on Modality, Normativity and Intentionality." *Philosophy and Phenomenological Research* 63: 611–623.

Wilson, E.O. (1975). *Sociobiology: The New Synthesis*. Cambridge: Belknap Press.

———. (1998). *Concilience: The Unity of Knowledge*. New York: Alfred A. Knopf.

Wright, Larry. (1973). "Functions." *Philosophical Review* 82: 139–168.